Naval Health Research Center

2

STRESS REACTIVITY: FIVE-FACTOR REPRESENTATION OF A PSYCHOBIOLOGICAL TYPOLOGY

R. R. Vickers, Jr.

Report No. 91-26

Approved for public release: distribution unlimited

NAVAL HEALTH RESEARCH CENTER
P.O. BOX 85122
SAN DIEGO, CALIFORNIA 92186-5122

NAVAL MEDICAL RESEARCH AND DEVELOPMENT COMMAND
BETHESDA, MARYLAND

Stress Reactivity:
Five-Factor Representation of a Psychobiological Typology[1]

Ross R. Vickers, Jr.

Cognitive Performance and Psychophysiology Department
Naval Health Research Center
P.O. Box 85122
San Diego, CA 92186-5122

[1]Report 91-26 was supported by the Office of Naval Research under Work Request ONR.WR.24030 and by the Navy Medical Research and Development Command, Bureau of Medicine and Surgery, Department of the Navy under Work Unit MR04101.00A-6004. The views presented are those of the authors and do not reflect the official policy of the Department of Navy, the Department of Defense, nor the U.S. Government.

Summary

Military personnel frequently must adapt to demanding, dangerous situations. Individual differences in personality are believed to play a significant part in the effect of such situations on performance and health, but the current understanding of how to best identify individuals who will adapt particularly well or poorly to demands is not very precise. The present study combined an emerging model of important general personality dimensions with prior work on stress responses in children and nonhuman primates to test the hypothesis that a typological classification could be developed which would identify individuals who can be expected to be resistant to stress and individuals who can be expected to be sensitive to stress.

Personality questionnaires completed by 3,328 male U.S. Navy recruits who volunteered to participate in studies of risk factors for infectious diseases provided the data for analysis. Analyses consisted of hierarchical and partitioning cluster analyses repeated in 10 random subsamples of the recruits to evaluate the replicability of clusters. The criteria for choosing the best clustering solution were whether similar clusters could be identified in all of the subsamples and how well each individual's cluster assignment in the analysis of data from his subsample could be predicted by applying classification formulas developed in the other subsamples to his personality profile.

Clusters with comparable group personality profiles could be identified in all 10 subsamples for 2-, 3-, 4-, and 5-cluster solutions, but not for a 6-cluster solution. The cluster assignments of individuals could be predicted with greater than chance frequency for the 2- through 5-cluster solutions. Each of these replicable solutions included one cluster which consisted of individuals who were above average on emotional stability, conscientiousness, extraversion, and agreeableness and a second cluster which had the opposite personality pattern. These two profiles were reasonably consistent with the a priori specification of the stress resistant and stress reactive types, respectively. However, the proportion of people in the resistant and reactive clusters was over 40% in the 2-cluster solution dropping to between 15% and 20% in the 5-cluster solution.

Subsets of individuals who possessed characteristics which are believed to represent stress reactivity/resistance in adult humans were reliably identified in these analyses. The five-cluster

solution was tentatively adopted as a measurement model for stress reactivity, because the estimated frequency of reactive and resistant types corresponded to a priori hypotheses and because this model includes groups that are not part of the hypothetical reactivity model. If further study shows that these unanticipated groups differ in important ways, the reactivity model would be partially disconfirmed. These results provide a simple method of classifying individuals as stress resistant or stress reactive that is grounded in a sound personality measurement model. Additional studies will be conducted to determine how well these personality types predict differences in stress reactions.

Introduction

Psychological stress is a commonplace element of everyday living which has significant implications for psychological and physical well-being. Most stress models assume that personality variables influence the person's reaction to stressful situations. There is less agreement about the specific personality constructs that must be considered to predict differences in reactions to stress. Attempts to delineate the critical psychological differences influencing stress reactions have produced an extensive body of poorly integrated findings that do not provide a coherent empirical basis for theory refinement. The present study attempted to combine a psychobiological model of stress reactivity arising from convergence of findings from studies ot children (Kagan, 1989; Plomin & Dunn, 1986) and nonhuman primates (Higley & Suomi, 1989; Sapolsky, 1990a, 1990b) with the five-factor model of personality (Digman, 1990; John, 1990). The objective of exploring the intersection of these two research models was to provide a general conceptual framework for defining stress reactive and stress resistant types in adult humans. Because the five-factor model can be used as a framework for classifying a wide range of individual difference measures, success in this endeavor was expected to provide a novel indicator of stress sensitivity in adult humans which could provide a basis for integrating some previously disparate findings on stress sensitivity.

General Structure of Stress Reactivity

As described by Higley and Suomi (1989), Sapolsky (1990a, 1990b), and Kagan (1989), stress reactivity has six important components. First, reactivity is indicated by a set of overlapping behavioral and endocrine indicators. The endocrine indicators represent integrated biological reactions to challenge that may influence short- and long-term adaptation. Second, individual differences in stress reactivity are stable over time, presumably because they represent genetic predispositions. Third, specific behavioral manifestations of reactivity depend on the social history of the individual, so the phenotypic expression of the genotypic pattern varies across individuals. Fourth, there is qualitative consistency in the phenotypic behavioral expression of reactivity within individuals over time. Fifth, differences in stress reactivity are observed primarily in response to acute stressors, particularly novel ones. Sixth, stress reactivity

is typological. In children, the estimated base rates for reactive and nonreactive types are 10% (Kagan, Reznick & Snidman, 1986) to 15% (Kagan, 1989).

The conceptual elements of stress reactivity are not new in personality research. Personality measurement models based on overlapping physiological and behavioral systems have been proposed (e.g., Eysenck, 1967, 1981; Strelau, 1983) which include genetic influences (Plomin, Chipuer, & Loehlin, 1990). Temporal and situational stability is the hallmark of personality constructs (Conley, 1984; Costa & McCrae, 1988), although the reactivity emphasis on the qualitative consistency of changing behavioral manifestations across time may differ from the usual conceptualization of stability. Modulation of genetic predispositions by social factors to produce variable phenotypic behavior patterns also has been suggested previously as a basis for typological models of individual differences (Meehl, 1973). Even the stress reactivity emphasis on integrating these topics to produce a viable model for normal personality has been anticipated (e.g., Eysenck, 1981). However, adopting the stress reactivity perspective can link these topics to another well-defined body of replicable findings that has been independently developed. A two-way exchange which can enrich both lines of study is possible.

<u>Stress Reactivity and the Five-Factor Model of Personality</u>

As noted above, stress reactivity is defined by a set of overlapping behavioral and endocrine indicators (Higley & Suomi, 1989; Kagan et al., 1986; Sapolsky, 1990a, 1990b). While behavioral indicators of stress reactivity in adult humans have not been precisely specified, there appears to be a consensus about the behavioral components of stress reactivity in the primate model. In a recent review summarizing observations relevant to that model, Higley and Suomi (1989) have described stress reactive animals as " . . . less likely to approach new stimuli, more anxious, more socially inhibited, and less likely to attempt challenging situations." The reactive animal also is described as being more acquiescent in social interactions and more likely to show depressive symptomatology when separated from other animals or surrogates it was reared with. Sapolsky (1990a, 1990b) provides a complementary description of a low reactivity animal as one who can correctly discriminate between threatening and non-threatening situations, who initiates direct aggression when threatened if he can win or who displaces aggression if he loses a fight.

The descriptions of reactive and nonreactive animals can be translated into personality constructs by applying the reactivity assumption that qualitative similarities are the key to understanding manifestations of reactivity in different organisms. With this assumption, the reactivity model can be linked to the five-factor representation of personality which is achieving increasing acceptance as a general representation of normal personality differences (Digman, 1990; John, 1990).

Brief definitions of the five dimensions comprising the five-factor representation of personality are needed to understand the potential links to stress reactivity. One dimension of the five-factor model, extraversion-introversion, contrasts socially outgoing individuals with relatively withdrawn individuals. A second dimension, neuroticism- emotional stability, contrasts individuals who experience strong negative affect, particularly under stress, with individuals who are emotionally stable. A third dimension, openness to experience, contrasts seeking out and enjoying new experiences with preferring familiar activities and situations. A fourth dimension, agreeableness-antagonism, contrasts conformity and efforts to get along with others with cynical, antagonistic attitudes and behaviors. The remaining dimension, conscientiousness-unreliability, contrasts methodical, achievement-oriented behaviors and striving for excellence with disorganized, unreliable behavior coupled with a willingness to accept lower standards of performance.

Table 1 is a tentative mapping of reactivity onto the five-factor model. This mapping was derived by considering the relationships between the five factor model and the descriptions of stress reactivity summarized above. Note that in this table, the two behavioral categories have been labelled "Stress Reactive" and "Stress Resistant." This labelling was chosen to emphasize the focus on both extremes of conceptualization. The term "stress reactivity" will be used below to refer to the overall conceptual model and the terms "reactive" and "resistant" will be used to refer to the specific subgroups when they are singled out.

Table 1

The Five-Factor Model and Stress Reactivity/Resistance

Big Five Dimension	Key Terms	Reactive	Resistant
Extraversion-Introversion	Socially-inhibited	Introverted	Extraverted
Neuroticism-Emotional Stability	Anxious, Depressed	Neurotic	Emotionally Stable
Openness to Experience-Inhibition	Dislikes new stimuli, inhibited	Inhibited	Open to Experience
Agreeableness-Antagonism	Acquiescent, aggressive	Agreeable	Antagonistic
Conscientiousness-Unreliability	Attempts challenges	Unreliable	Conscientious

The mapping of stress reactivity onto the five-factor model of personality is reasonably straightforward except in the case of conscientiousness. In this case, the key consideration is the link between conscientiousness and setting high standards and striving to perform well. This element of conscientiousness is akin to achievement motivation and even has been referred to as "will to achieve" (Digman & Takemoto-Chock, 1981). The achievement motivation literature indicates that this type of motivation is associated with a willingness to take on challenging performance requirements (Atkinson & Feather, 1966). The proposed mapping of conscientiousness onto the reactivity element of attempting challenging situations is based on the potential equivalence of these two sets of descriptions.

The proposed mapping of stress reactivity onto the five-factor personality model is tentative. This mapping should be regarded as representing a set of related hypotheses as other plausible mappings probably could be developed. At this point in research on the stress reactivity model in adult humans, the proposed mapping defines expected patterns of personality attributes for reactive and resistant types and illustrates the potential for using the five-factor model to develop a suitable measurement model for generalizing the stress reactivity model from

nonhuman primates and children to the general human population. One perspective on the analyses conducted in this study is that they were designed to test this set of hypotheses rather than adopting the model by fiat.

Reactivity as a Typology

Table 1 implicitly differs from the usual application of the five-factor model of personality and stress reactivity by describing personality differences in terms of two distinct types rather than five continua. The five-factor model has been developed from a research tradition that regards continuous dimensions as the appropriate basis for characterizing individual differences in personality. Stress reactivity assumes the existence of a typology, rather than a set of continua. While typologies are a recurrent topic in personality research (e.g., Gangestad & Snyder, 1985; Jemmott, et al., 1990; Myers, 1980; Strube, 1989), no available typology can be confidently equated with stress reactivity. A suitable basis for stress reactivity classifications must be developed before meaningful comparisons to these other constructs are possible. This aspect of the work presents special conceptual and analytic problems. Even though some attempts have been made to combine typological and continuous constructs conceptually (e.g., Hogan, 1983), it is more common to treat the two as alternative models to be contrasted (e.g., McCrae & Costa, 1989; Snyder & Gangestad, 1986; Strube, 1989).

The present attempt to develop a typological classification involves mapping five behavioral dimensions, each with evidence of discriminant validity in descriptions of humans (e.g., McCrae, 1982; McCrae & Costa, 1987), onto three behavioral types. At present, claims for a typology appear to rest heavily on the temporal stability of classification into high and low reactive categories (Kagan, 1989; Kagan et al., 1989; Kagan, Reznick & Snidman, 1986; Kagan, Reznick, Snidman, Gibbons & Johnson, 1988). The fact that only some individuals are consistently classified as falling in these extreme groups over repeated measures is taken as evidence that reactivity is typological.

The pattern of results that has been the basis for inferring the existence of stress reactivity types could be obtained even if differences in reactivity truly represented a continuous dimension. In this case, measurement errors would be expected to produce variability in determinations of whether a person or animal fell above or below the cutoff at different times. Individuals who were extreme on a reactivity continuum measured with error would be expected to remain above

or below the median for the sample because they would have to have exceptionally large errors in measurement to shift their scores to the other side of this criterion point in the distribution. Therefore, the simple demonstration that subtypes of individuals showing the patterns of personality scores hypothesized in Table 1 can be identified in the population would be one step toward confirming the existence of a reactivity typology.

Criteria for Detecting a Typology

The present study employed cluster analysis to test for the existence of reactive and resistant types using five-factor personality measures. The establishment of a priori criteria was important, because clustering algorithms are designed to produce clusters from virtually any data. As a result, determining the appropriate number of clusters to extract from a data set is a major problem for cluster analyses (Blashfield & Aldenderfer, 1988; Milligan & Cooper, 1987). In an attempt to address this problem in an orderly fashion, three decision criteria for evaluating alternative solutions were established prior to undertaking the analysis.

The first criterion was that the types or clusters should be consistently identified in the population. Operationally, this criterion was applied by dividing a large sample into random subsamples and examining the consistency with which comparable clusters were identified across the subsamples.

The second criterion was that the extent of clustering should exceed that expected by chance. Operationally, this criterion was applied by utilizing statistics developed to describe the extent of clustering within a sample which contrast the observed degree of clustering with the level expected by chance. The level expected by chance was determined from purely statistical models and from prior Monte Carlo research on clustering.

The third criterion was the closeness of the match between the observed typology and the theoretical reactivity typology. This last criterion was operationalized in terms of two distinct comparisons. First, the observed personality profiles for clusters were compared to the hypothetical reactivity/resistance profiles in Table 1. Second, the base rates of occurrence for the groups which provided the best matches to these hypothetical profiles were compared to the estimated 10% to 15% reported by Kagan and his colleagues (e.g., Kagan, 1989).

Method

Sample

Study participants (n = 33[illegible]) were male U.S. Navy recruits who volunteered to participate in studies of the influence of psychological factors on susceptibility to disease during recruit training. The typical study participant was 19.7 years of age (S.D. = 2.76; range = 16-34). Whites (70.2%) were the predominant ethnic group with Blacks (15.3%) and Hispanics (7.3%) the primary minority groups. Most recruits had completed 12 years of schooling (67.4%), but a sizable minority had more schooling (24.8%), and only a few had less than a high school education (7.8%).

Personality Measures

The NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985) provided standardized measures of the five-factor personality model. Although the NEO-PI includes facet scales assessing 18 specific personality traits comprising 3 of the 5 major dimensions, the present analyses were conducted using measures of just the five major dimensions comprising the five-factor model. This decision was based on a desire to use the simplest general personality model that could reasonably accomplish the goals of the study. Scale scores were computed by taking the mean of the responses to individual items assigned to each dimension. This approach was adopted in preference to the standard practice of employing the sum of the item responses because the summing procedure would increase the variance of the longer scales relative to the shorter scales. When the clustering algorithms were used, the longer scales would then be given more effective weight in defining the similarity between different individuals (see Analysis Procedures below). Scores were computed for each dimension provided that no more than 10% of the items comprising that dimension were missing from the questionnaire responses.

Analysis Procedures

Overview. The analysis procedures were designed to evaluate the typology claim by determining whether cluster analyses reliably identified groups of individuals with similar personality profiles when the procedures were applied to independent samples. The assumption was that if different types truly existed in the population being studied, any reasonably large sample from that population would contain enough of that type to be identified in cluster analyses. Furthermore, the consistency with which clusters could be replicated when different

numbers of clusters were extracted from the data was assumed to provide a useful index of the true number of clusters present in the population. The assumption that different types existed in the population was tested by performing a series of hierarchical agglomerative cluster analyses in different subsamples to define possible types. The resulting clusters then were matched across subsamples to determine how replicable the clusters were at each level of clustering. The reliability of assignment of individuals to particular clusters then was checked by comparing their cluster assignment from the analysis of data in their own subsample to their cluster assignment based on jackknifed classification functions from the other subsamples. These steps then were repeated for a partitioning cluster analysis procedure to determine how the choice of clustering procedures affected the results and to develop a final basis for classifying individuals. Details of these procedures and the statistics used to compare the groupings provided by solutions with different numbers of groups are given below.

Subsample Definition. The sample was divided into 10 subsamples by assigning each individual to a group based on the last digit of his social security number. This procedure was employed to keep the number of participants per analysis small enough to run the clustering programs efficiently and to provide enough samples to evaluate the replicability of clusters across samples of moderate size. Subsample size ranged from 331 participants to 335 participants due to small amounts of missing data.

Development of Initial Clusters. The initial assignment of individuals to clusters was accomplished by hierarchical agglomerative clustering analysis of individuals within each subsample. These analyses were conducted with the SPSS-X program "CLUSTER" (SPSS, Inc., 1988) with Ward's (1963) method chosen as the clustering procedure based on performance in simulation studies (Blashfield, 1976; Blashfield & Aldenderfer, 1988; Milligan, 1981b; Milligan & Cooper, 1986, 1987). Ward's method was applied with squared Euclidean distance as the similarity measure. This distance measure was computed using each individual's scores on the five primary dimensions of the NEO Personality Inventory. Each individual's group assignment was determined for the 2- through 6-group levels of clustering, and his assignment for each level was saved for use as a group classification variable in later analyses. The clusters defined in this stage of the analysis defined groups of individuals within each subsample and are referred to as "basic clusters" in the following discussion.

Constructing Centroid Data Files. The second step in the analyses constructed data files describing the personality profiles of the clusters identified in the hierarchical agglomerative analysis of data from individual participants. This step was a precursor to analyses to match clusters across subsamples as described below. The average personality profile (i.e., the group centroid) was determined for each cluster defined in the preceding analysis phase. Thus, a pair of group profiles were generated to represent the two groups defined when only two clusters were extracted, three profiles were generated to represent the three groups when three clusters were extracted, and so on. Separate profiles were computed for each level of clustering within each subsample. These group centroids then were used to create five separate data files, one for each level of clustering. For example, one file consisted of 20 observations, each observation representing the group centroid for one of the two groups defined in the 2-cluster solution in each of the 10 subsamples. Similar files with 30, 40, 50, and 60 observations were constructed for the 3-cluster, 4-cluster, 5-cluster, and 6-cluster analyses.

Clustering Centroids. Data files comprised of cluster centroids were constructed because it could not be assumed that the group assigned the label "Cluster 1" in one subsample was necessarily equivalent to the group with the same label in another subsample. To match clusters across subsamples, the group centroid data were analyzed by the same hierarchical agglomerative procedures applied to the data from individuals, except that the number of clusters extracted was equal to the number of clusters defined in the analysis giving rise to the centroid data set (e.g., 2 groups for the 2-cluster data set). The resulting clusters are referred to below as "higher-order clusters" to indicate that they were clusters of the basic level clusters.

Discriminant Analysis of Centroids. Following the definition of higher-order clusters, a discriminant function analysis was performed with the group centroids as observations and higher-order cluster assignment as the grouping variable. The probabilities of membership for each basic cluster in each higher-order cluster were computed from this discriminant function analysis. These probabilities then were used to match basic level clusters across samples. Matching was initiated by assigning each basic cluster to that higher-order cluster for which it had the highest probability of membership. In most cases, assignment resulted in one basic level cluster from each subsample being assigned to each of the higher-order clusters. When two basic level clusters from a subsample were assigned to the same higher-order cluster in this first step,

step, the basic cluster with the highest probability of membership was retained as the best example of that higher-order cluster in that subsample. The remaining group was assigned to the higher-order cluster that did not have a match in that subsample in the initial analysis.

Discriminant Analysis of Data for Individual Participants. The process of matching clusters across subsamples provided a tentative assignment of individuals to groups. It then was necessary to know how the personality profiles of these groups compared to the hypothesized personality profile in Table 1. It also was desirable to determine the reliability with which individual participants could be assigned to particular groups. This latter point was of interest as a possible basis for deciding the appropriate number of groups to extract to represent the population. Discriminant analyses of the data for individual study participants were conducted to answer these questions. The group classification variables in these analyses were based on the matching of groups in the previous step. Separate discriminant functions were performed in each subsample for the 2- through 5-cluster groupings derived in the initial clustering of individuals. The 6-cluster grouping was not included in this phase of the analysis because it proved to be unreliable across subsamples (see Results). These discriminant analyses produced the classification function coefficients used in the next step of the analysis.

Jackknifed Predictions of Group Membership. Whenever possible, it is desirable to have an external criterion to evaluate the precision with which clustering procedures have recovered true group membership. In the present analyses, no such external criterion was available. For this reason, a jackknifed criterion for assigning individuals to groups based on the personality profiles was constructed. This criterion combined the results from 9 other subsamples to predict cluster membership for individuals in the 10th subsample. The convergence of the predictions and the actual cluster assignment provided a summary measure of the replication of group assignments. The present procedure was adopted in preference to a series of pairwise comparisons between samples to reduce the computational load and to capitalize on the effects of aggregation as a means of stabilizing estimates of the classification function coefficients.

Jackknifed classification functions were produced by averaging the classification function weights for 9 samples with the 10th sample held out. The average classification weights then were applied to predict group membership for the individuals in the subsample which had been held out. This procedure was repeated for each subsample for each level of clustering from 2

through 5 groups. Tables then were constructed to determine the match between the individual's initial group assignment and his jackknifed classification. Cluster membership defined in the initial hierarchical analysis of the data for that subsample was the row variable in these tables. Cluster membership based on the jackknifed classification functions was the column variable. The tables thus generated were used to compute the kappa and Rand statistics described below.

Statistical Tests for Clustering. The statistics used to assess the extent of clustering in the data included the Rand statistic (1971) with the Hubert and Arabie (1985) adjustment for chance correspondence and Cohen's (1968) kappa statistic. The Rand statistic employs a more lenient definition of consistent classification than does kappa. A pair of observations is considered to be consistently classified if both are assigned to the same groups in both cluster solutions or if both are assigned to different groups in both solutions. The first type of consistent pair would be represented by individuals falling in the same cell within a crosstabulation. The second type would consist of individuals who were in different rows and different columns within a crosstabulation. By a process of elimination, therefore, inconsistent classifications include any pairs of individuals who were in the same cluster on the basis of one classification, but different clusters on the basis of the other classification. These individuals would be either in the same row, but different columns, or the same column, but different rows, in the crosstabulation. The Rand statistic, therefore, is

$$\text{Rand} = 1 - \frac{\text{Total Pairs} - \text{Pairs in the Same Cell}}{\text{Total Pairs}}$$

The Hubert-Arabie adjustment to the Rand statistic corrects for the number of consistent classifications expected by chance given the marginal frequencies. With this adjustment, the Rand statistic, like kappa, measured the extent to which the observed matching of subsample and jackknifed classifications exceeded that expected by chance.

The kappa statistic is computed with the assumption that groups from the two variables defining a cross-tabulation can be mapped in a one-to-one fashion. Thus, a match would occur in a present instance if a person's assignment in his subsample analysis and the jackknife analysis were to the same group. Formally,

$$\text{Kappa} = \frac{\text{Correct Matches - Matches Expected by Chance}}{\text{Total Pairs - Matches Expected by Chance}}$$

where the number of matches expected by chance is the marginal probability for the row (representing the matched group in one classification) times the marginal probability for the column (representing the matched group in the other classification) times the total number of cases.

The Rand and kappa statistics differ with respect to their dependence on the higher-order clustering procedure. Kappa values will be affected by this procedure because they depend on which group is designated as representing a given type in that clustering. In contrast, the Rand statistic does not depend on matching between specific groups. Instead, it is based on the pattern of cell frequencies over the entire cross-tabulation and is invariant over any transposition of rows and/or columns. This statistic therefore is independent of the higher-order clustering except insofar as clustering influences the computation of the classification functions. Note, however, that the classification function computations applied to any given subsample are completely independent of the data for that subsample. Thus, the Rand statistic depends on classification functions based entirely on the data from other subsamples and on the clustering analysis produced by the current subsample. These two classifications are completely independent except insofar as a participant's assignment in both classifications is derived from his personality profile. Because of this independence and the availability of Monte Carlo evaluations of the expected values of the Rand statistic, it provides a suitable criterion for the presence of clustering which is not based on any assumptions about the matching of groups across subsamples. Kappa complements this criterion by showing the increment in predictive accuracy over chance that is achieved if one assumes a one-to-one mapping of groups from the subsample into the higher-order classification.

Partitioning Cluster Analyses. Additional analyses were undertaken to develop clusters based on the partitioning routine provided by the SPSS-X QUICK CLUSTER procedure. Partitioning analyses have provided the most effective recovery of true group membership in simulation studies if appropriate initial estimates of the group centroids are available (Milligan

& Cooper, 1987). Simulation studies also have indicated that centroids from hierarchical clustering procedures provide suitable starting points for effective partitioning analyses.

In the present analyses, oneway analyses of variance of the data from individual participants were performed to determine the centroids to be used as starting points for the partitioning analysis. In these oneway analyses of variance, each participant's group membership was determined by the higher-order cluster to which his basic level cluster was assigned in the preceding hierarchical analyses. The basic level cluster to which the individual was assigned was determined from the initial hierarchical cluster analysis. It is important to emphasize that although the jackknifed classifications were treated as a relatively objective external criterion when evaluating the accuracy of cluster assignments in the hierarchical analysis, the jackknifed classifications were not used as the basis for group assignment when determining starting values for the partitioning analysis. The steps for the hierarchical analysis then were repeated for the partitioning analysis culminating in another contrast of subsample cluster assignment with jackknifed cluster assignment.

Criteria for Evaluating Cluster Solutions. The criteria for comparing the alternative cluster solutions included:

(a) 1:1 Mapping: The number of subsamples which produced a cluster representing each of the corresponding set of higher-order clusters. This criterion was a test of the fundamental assumption that if different types truly exist in the population, they should be identifiable in most or all analyses of samples drawn from the population. In the present case, the extent to which a 1:1 mapping was present was determined on the basis of the initial assignments of basic clusters to higher-order clusters (see Discriminant Analysis of Centroids above).

(b) Minimum Cluster Size: This figure represents the number of clusters assigned to each higher-level cluster on the basis of the initial assignments of basic clusters to higher-order clusters (see Discriminant Analysis of Centroids above). If each sample produced one example of each higher-order cluster, the minimum cluster size would be 10. The reported minimum cluster size complements the 1:1 Mapping criterion by indicating recovery of higher-level clusters at the level of individual types in terms of the poorest recovery of any one type at that level of clustering.

(c) Consistency of Proportions: An estimate of the stability of frequency of individuals in the higher-order clusters, indicated by the chi-square/degrees-of-freedom ratio for the crosstabulation of higher-order cluster membership by subsample.

(d) Kappa: The average and standard error for the kappa statistics computed in the subsamples.

(e) Rand: The average and standard error for the Rand statistics in the subsamples plus a t-test comparing these observed values to the expected value based on Monte Carlo analysis of random data.

(f) Adjusted Rand: The average and standard error for the Hubert-Arabic adjusted value of the Rand statistic. A t-test was computed with the null hypothesis that the Adjusted Rand = .10. This null value was slightly higher than the upper limit of the reported 95% confidence interval for the Adjusted Rand statistic derived from Milligan's (1981a) Monte Carlo analyses. This nonzero null hypothesis was chosen to provide a relatively stringent null hypothesis when testing for the presence of clustering.

Results

Hierarchical Agglomerative Clustering

Between 2 and 5 clusters could be identified reliably in the hierarchical clustering, but attempts to identify a sixth cluster were unsuccessful (Table 2). The major basis for this conclusion regarding the 6-cluster solution was that only 3 of 60 basic clusters were assigned to the sixth group when 6 clusters were extracted in the higher-order clustering analyses. This figure compared to a minimum of 6 groups assigned to each higher-order cluster when 2 to 5 clusters were extracted. If an example of a particular cluster could not be identified in at least half of the subsamples, it was considered reasonable to conclude that this type was not reliably identified.

Considering the solutions for 2 to 5 clusters, one important finding was that the statistics used to estimate the extent of clustering were well above what would be expected by chance for all of the solutions. When attention is directed toward selecting the appropriate level of clustering, the results of the hierarchical analysis were equivocal. Either the 2-, 3- or 5-cluster solutions would be favored depending on the criterion chosen. The 2-cluster solution produced the best kappa criterion value. The three-cluster solution was the most complex solution that produced examples of each cluster in all 10 subsamples and was the theoretically appropriate solution, but was not optimal by any other criterion. The five-cluster solution produced the

Table 2
Summary of Cluster Analysis Results

	2-Group	3-Group	4-Group	5-Group	6-Group
Hierarchical Analysis					
Subsamples with 1:1 Mapping	10	10	7	5	2
Minimum Cluster Size	10	10	6	7	3
Consistency of Proportions*	19.99	27.25	20.97	13.56	
kappa					
Mean for Subsample	.548	.368	.417	.323	
Standard Error	.073	.059	.040	.050	
Rand					
Mean for Subsamples	.674	.682	.714	.796	
Standard Error	.033	.035	.011	.019	
t-test**	3.15	3.20	13.09	11.89	
Significance	.006	.006	.001	.001	
Adjusted Rand					
Mean for Subsamples	.339	.358	.365	.495	
Standard Error	.069	.069	.019	.060	
t-test***	3.46	3.74	13.95	6.58	
Significance	.004	.003	.001	.001	
Partitioning Analysis					
Subsamples with 1:1 Mapping	10	10	10	10	6
Minimum Cluster Size	10	10	10	10	8
Consistency of Proportions*	7.00	5.98	4.39	2.99	
kappa					
Mean for Subsamples	.870	.722	.700	.716	
Standard Error	.022	.022	.030	.025	
Rand					
Mean for Subsamples	.872	.767	.819	.837	
Standard Error	.017	.015	.012	.011	
t-test**	17.76	13.13	20.75	24.27	
Significance	.001	.001	.001	.001	
Adjusted Rand					
Mean for Subsamples	.743	.508	.548	.527	
Standard Error	.033	.028	.029	.031	
t-test***	19.48	14.57	12.00	13.77	
Significance	.001	.001	.001	.001	

* Chi-square/degrees-of-freedom ratio for test of consistency of proportions for cluster membership across subsamples.

**t-test (9 df) for the null hypothesis that Rand = .57 based on Milligan's (1981a) Monte Carlo results for random data.

***t-test (9 df) for the null hypothesis that Adjusted Rand = .10, an upper limit for the estimated value for random data based on Milligan's (1981a) Monte Carlo analyses.

largest Rand statistic and Adjusted Rand statistic. Because none of the solutions could be conclusively ruled out of consideration, all four levels of clustering were retained for the partitioning analysis.

Partitioning Analysis

The partitioning analysis produced more consistent results than the hierarchical analysis. Each subsample produced an instance of each higher-order cluster up through the analysis involving 5 clusters. The proportion of individuals assigned to each cluster was more consistent across subsamples than in the hierarchical analysis, as indicated by smaller chi-square/degrees of freedom ratios. The kappa and Rand statistics substantially exceeded what would be expected by chance for all levels of clustering, but were consistently largest for the 2-cluster solution. The remaining three solutions produced approximately comparable values for the kappa and Rand statistics with the 3-cluster solution producing the largest kappa, the 4-cluster solution producing the largest Adjusted Rand, and the 5-cluster solution producing the largest Rand value. Given the jackknifed estimates of the sampling variability of these statistics and allowance for multiple comparisons (Dunn, 1958), it is reasonable to assert that the differences between the 3-, 4-, and 5-cluster clustering measures probably did not exceed chance.

Classification Probabilities, Sensitivity, and Specificity

The 3-cluster and 5-cluster solutions were considered further despite the fact that the 2-cluster solution produced the largest kappa and Rand values in the partitioning analysis. This decision was made because the 3-cluster solution was theoretically appropriate, while the 5-cluster solution provided the highest level of reliable cluster differentiation in the data that was robust across clustering methods. In addition, it was judged that extracting too many groups was a less serious decision error than extracting too few groups. The reasoning was that the clusters defined at different levels are approximately hierarchically related because hierarchical agglomerative clustering was employed as the first analysis step. If a 2-cluster solution were correct in the sense that there really were only two distinct types present in the population, the more complex cluster solutions would essentially be creating artificial subtypes within true types. Any comparisons between the subtypes with regard to behavioral or biological attributes which would be pursued to validate the model would show that the subtypes did not differ significantly. Given that evidence, a retreat to a conceptual model based on two types would be possible.

However, if more than two types are present and a model based on two types were adopted initially, the analyses which would be expected for typical comparisons between the types will not necessarily produce clear evidence of heterogeneity within types to indicate the need for further differentiation. The consequences of overextraction should be readily evident in any systematic attempts to validate the typology, but the consequences of underextraction would not necessarily be evident. For this reason, overextraction was a more acceptable error than underextraction at this phase in the research.

Treating the jackknifed classification as an approximation to an objective classification, the sensitivity and specificity of the 3- and 5-cluster solutions can be compared (Table 3). Sensitivity is the proportion of true examples of a given type correctly identified by one's test. Operationally, this statistic was computed as the proportion of individuals classified into a cluster in the jackknife assignment (the "true" criterion) who also were classified into that group in the subsample cluster analysis. The 5-cluster classification had a lower median sensitivity (.834) than the 3-cluster classification (.895).

Table 3
Jackknife Cross-Validation of Classifications

Subsample Group	Jackknifed Group: 1	2	3	4	5	Per Cent	Sens.	Spec.
Three-Group								
1	1251	97	42			49.9%	.753	.900
2	219	824	0			27.7%	.895	.790
3	191	0	704			22.4%	.941	.787
%	41.8%	31.3%	26.9%					
Five-Group								
1	766	41	44	37	1	26.7%	.730	.862
2	92	531	46	17	25	21.4%	.834	.747
3	82	29	504	0	28	16.5%	.755	.784
4	109	21	0	419	0	16.1%	.886	.763
5	1	15	74	0	446	19.3%	.882	.832
%	31.6%	19.1%	14.2%	15.0%	20.1%			

NOTE: "Sens" = Sensitivity. "Spec" = Specificity. Definitions are given in the text. "%" = Percentage of total sample assigned to the group in the jackknifed classification.

Specificity is the proportion of observations classified as a particular type by the clustering algorithm that are examples of that type when classification is objectively determined. Operationally, this statistic was computed by determining the proportion of cases predicted to be in a given cluster on the basis of the subsample analysis who were assigned to that cluster in the jackknife analysis. The median specificity for the 3-cluster analysis (.790) was slightly higher than the comparable value for the 5-cluster analysis (.784).

Relationship between 3-Group Solution and 5-Group Solution

Although both the sensitivity and specificity of the 3-cluster solution was better than that of the 5-cluster solution, several considerations led to the retention of the 5-cluster solution for subsequent analyses. One consideration was that the overall gain in the convergence between subsample-based cluster assignment and jackknifed classifications achieved by adopting the 3-cluster solution was only 113 matches (3.4%). This gain was trivial, particularly in light of the greater range of errors which would be expected by chance given the two additional groups in the five-cluster solution.

A second consideration was that the key groups in the 3-cluster solution could be readily identified in the 5-cluster solution. Based on the mean profiles for the groups in the 3-cluster solution, clusters 2 and 3 were candidates for reactive and resistant groups. The fourth group in the 5-cluster solution was almost entirely a subset of the second group in the 3-cluster solution (472 of 473). The fifth group in the 5-cluster solution was comparably similar to the third group in the 3-cluster solution (500 of 501). Thus, any differences that would be obtained between the two candidate reactivity/resistance groups in the 3-cluster solution also should be evident in the 5-cluster comparisons. However, the 5-cluster classification produced subgroups which represented about 15% of the total sample, thereby satisfying an a priori criterion for choosing between alternative solutions.

The final consideration in adopting the five-cluster solution was that by adding groups beyond those hypothesized in the reactivity model it would be possible to falsify this model in later analyses. Falsification could be achieved by developing evidence that distinguishing the additional groups from the reactivity groups provided improved prediction of reactivity-relevant criteria relative to what would be achieved with only three groups. There appeared to be no major cost associated with retaining enough groups to permit model falsification because the

reasonably clean hierarchihcal structure of the typology provided a simple basis for redefining groups and selecting an alternative level of clustering if further study showed this to be appropriate.

Personality Profiles for the Clusters

Figure 1 presents group profiles for the 5-cluster classification; means and standard deviations for the groups are given in the appendix which also includes the classification functions for any readers interested in applying the typology to other data. This figure is intended to convey several important points. First, there were two extreme groups with basically opposite profiles Neuroticism and Conscientiousness were the two dimensions that differed most between these groups. These dimensions were placed at opposite ends of the profile to emphasize that Neuroticism is high and Conscientiousness low in one extreme group while the other shows the reverse pattern.

The second point is that the remaining three groups had profiles that were near the sample mean compared to the two extreme groups. This point is evident if one compares the range of mean values for these groups to the range for the extreme groups.

The third point illustrated by Figure 1 is that the variables which differentiate between the intermediate groups are Openness, Agreeableness, and Extraversion. Two of the groups were very near the sample mean for Neuroticism and Conscientiousness, but either above the mean on the remaining three dimensions or below the mean on those dimensions. The final group was essentially average with regard to personality dimensions except for an above average score on Conscientiousness.

A fourth point illustrated by Figure 1 is that the personality dimensions differed in how well they discriminated between groups. If the highest and lowest mean scores in Figure 1 are identified for each dimension, the other four dimensions show a much wider range of differences than does Openness to Experience. Openness, therefore, was of relatively limited importance in discriminating between the clusters.

A final critical point that is illustrated by Figure 1 is that none of the profiles corresponded precisely to the hypothetical reactivity or resistance profiles. If the predicted pattern of

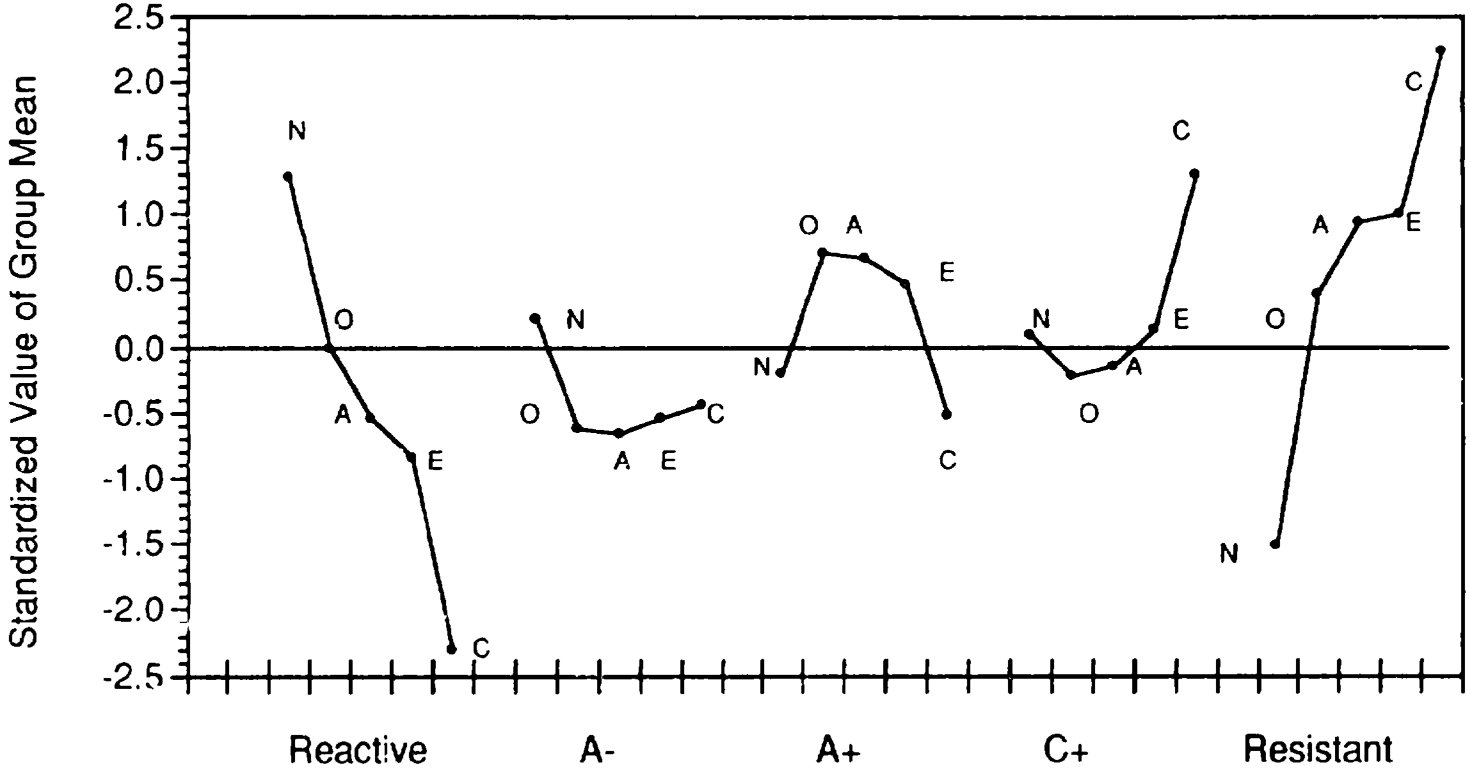

Figure 1. Personality Profile for Five-Group Typology

* Note: N = Neuroticism; O = Openness; A = Agreeableness; E = Extroversion; C = Conscientiousness. See text for definition of groups. Standardized means computed using pooled within-group standard deviation.

differences in Table 1 were plotted in Figure 1, the result would resemble an "M" for reactive individuals and a "W" for resistant individuals. None of the observed profiles match these hypothesized profiles.

Discriminant Function Analysis. A more compact representation of the cluster differences is provided by considering three canonical discriminant functions which characterized the personality differences between clusters. Three functions were used to summarize these differences because analyses indicated three statistically significant discriminant functions in 7 of 10 subsamples, two significant discriminant functions in one subsample, and four significant discriminant functions in the remaining two subsamples.

Average discriminant function weights were computed by matching functions across subsamples on the basis of which personality scale had the largest standardized loading on the function. When rotated, two functions corresponded to the basic NEO-PI measures of Conscientiousness and Neuroticism (Table 4). The third discriminant function was defined

Table 4

Standardized Discriminant Functions

Scale	Function 1 Mean	Function 1 S.D.	Function 2 Mean	Function 2 S.D.	Function 3 Mean	Function 3 S.D.
Conscientiousness	.991***	.021	.014	.067	.009	.066
Agreeableness	.032	.101	.608*	.300	-.059	.220
Openness to Experience	.174*	.087	.309	.284	-.139	.331
Neuroticism	-.061	.170	-.046	.236	.884**	.125
Extraversion	-.080	.088	.538*	.240	-.053	.346

NOTE: Table values are the mean and standard deviations for the standardized discriminant function weights across the 10 subsamples.

* $t > 1.83$, 9 df, $p < .05$, one-tailed

*** $t > 4.30$, 9 df, $p < .001$, one-tailed

primarily by higher scores on Agreeableness and Extraversion. Openness to Experience had relatively small loadings (<.31) on all three functions. Despite the significant t-test for the loading of Openness on the first discriminant function, there did not appear to be a clear basis to uniquely link Openness to Experience to any one of the three discriminant functions. For this reason, the discriminant functions will be referred to as the Conscientiousness function, the Neuroticism function, and the Agreeableness/Extraversion function.

Using .40 (absolute) as the criterion to define moderate deviations from average and .70 (absolute) to define large deviations from average, the profiles can be summarized as follows:

a) The "React[illegible] [illegible]p in Figure 1 combined a large negative deviation on the Conscientiou[illegible]n (z = -2.46) with a large positive deviation on the Neuroticism fu[illegible]n (z = 1.26) and a moderate negative deviation on the Agreeableness/Extraversion function (z = -.89).

b) The "Resistant" group in Figure 1 combined a large positive deviation on the Conscientiousness function (z = 2.40) and on the Agreeableness/Extraversion function (z = 1.48) with a large negative deviation on the Neuroticism function (z = -1.41)

c) The remaining groups produced means near zero (< .53 absolute) on 2 of 3 canonical functions, but "A-" was moderately below average on the Agreeableness/Extraversion function (z = -1.04), "A+" was moderately above average on this function (z = 1.05), and "C+" was well above average on the Conscientiousness function (z = 1.33).

Correlations between Personality Measures within Groups. Because Meehl (1973) has suggested that typology indicators will be uncorrelated within types, the within-group correlations between personality dimensions were examined. The analysis examined the variability of correlations across clusters to determine whether a single set of averaged correlations could be used to represent the 5 within-cluster correlation matrices. The variation in correlations across clusters was nominally significant except for the Openness-Conscientiousness correlation (Table 5), but the statistical significance of the variability in correlations across the subgroups was largely attributable to sample size. Hoelter's (1983) critical N, also given in Table 5, indicates the maximum sample size which would have produced a nonsignificant chi-square given the observed variability of the correlations across subgroups. Hoelter (1983) has recommended a critical N of 200 times the number of groups compared as a criterion for accepting a model as

fitting the data. In the present instance, this criterion would lead to the acceptance of the null hypothesis of equal within-group correlations for 9 of 10 correlations if $p < .05$ were the significance criterion. A Bonferroni adjustment to allow for the number of significance tests performed would yield $p < .005$ as a more appropriate criterion. Applying this criterion would have produced critical Ns 1.57 times larger than those in the table, ranging from 1539 to 11,500.

Table 5

Average Within-Group Correlations between Personality Measures

		Average r	Chi-Square	Critical N
Neuroticism with:	Extraversion	-.004	23.70	1332
	Openness to Experience	.110	25.35	1245
	Conscientiousness	.024	14.72	2145
	Agreeableness	-.001	14.34	2201
Extraversion with:	Openness to Experience	.138	32.13	983
	Conscientiousness	-.075	18.80	1679
	Agreeableness	.018	21.02	1502
Openness to Experience with:				
	Conscientiousness	-.051	4.30	7343
	Agreeableness	-.027	17.86	1767
Conscientiousness with:	Agreeableness	-.004	10.82	2918

NOTE: The average correlation is the weighted average of the within-group correlations for the five-group classification. The chi-square indicates the variability of the observed correlations about the average (Cohen & Cohen, 1983); $p < .05$ if chi-square > 9.49 (4 df). The critical N is the sample size that would produce a nonsignificant chi-square (Hoelter, 1983).

With appropriate allowance for sample size and the number of significance tests performed, the average correlations were representative of the within-cluster correlations. The largest absolute value was .138, and only 2 of 10 correlations were as large as .10 absolute. Openness to Experience generally produced larger within group correlations (average absolute

$\underline{r}$ = .082; range = .027 - .138) than the other variables (average absolute $\underline{r}$ = .021; range = .004 - .075). In fact, only one correlation that did not involve Openness to Experience was as large as the smallest correlation that did involve Openness. At a minimum, therefore, the correlations between Neuroticism, Extraversion, Conscientiousness, and Agreeableness were effectively equal to zero.

DISCUSSION

Stress reactivity is a viable typology for personality differences in U.S. Navy recruits. Two subgroups which were reasonable approximations to the hypothetical stress reactive and stress resistant types described in the introduction consistently were identified when self-descriptions were analyzed. The analysis procedures were selected to impose the minimum number of preconceptions as constraints on the determination of the cluster solutions and to ensure that the final cluster solution provided above chance evidence of clustering that was replicable across subsamples of respondents. Given the nature of the analysis procedures, it was quite possible that the results would have indicated that no reliable clusters were present in the data or that reliable clusters were present which bore no similarity at all to the hypothesized reactivity/resistance groups. In this context, therefore, the identification of replicable clusters which could be reasonably equated with reactivity and resistance was encouraging.

The empirical profiles for the presumed reactive and resistant groups had three major components which could interact to heighten or minimize (respectively) the effects of exposure to demanding environments, particularly the social elements of those environments. One component was a general tendency toward negative emotional states, particularly depression and anxiety, which the person reports is exacerbated by stress. The remaining attributes involve patterns of social interaction that could generate stress. One social interaction component was the tendency to be disagreeable, cynical, and socially withdrawn. The second social interaction component was the inability to formulate and follow through on courses of action in task related activities and to be generally unreliable. A cynical, withdrawn individual obviously possesses attributes that can generate interpersonal friction through his or her negative expectations about others. When challenges are encountered, these attributes would make it difficult to benefit from

the support that social interactions can provide because of general isolation from the social group. The unreliability of the reactive individual could generate stressful conditions by failures in the work arena or by contributing to day-to-day problems such as not remembering to pay bills on time, forgetting appointments until the last minute, and so on. These types of failures are likely to elicit negative responses from the social environment and may contribute further to feelings of isolation and cynicism. The stress resistant personality pattern is the opposite of the gloomy sketch for stress reactive individuals.

The proposed interpretation of the reactive and resistant personality profiles suggests a number of possibilities regarding the specific mechanisms by which stress reactivity/stress resistance influence the outcome of exposure to a stressful situation. However, at this time, these profiles must be regarded as provisional, and the associated inferences regarding psychosocial dynamics clearly are speculative.

The value of speculation about the psychosocial dynamics of stress resistance depends heavily on the legitimacy of the claim that distinctive types can be identified in the population. In this regard, the fact that reactive and resistant types can be reliably identified in different samples is nontrivial. Even though clustering algorithms are designed so that they will produce clusters in any data, empirically-derived clusters would be expected to differ from one sample to the next if no real clusters were present in the population. The fact that the present study demonstrated that the same clusters could be consistently recovered from self-reports suggests that real clusters were present in the population. The accuracy of subsample classifications relative to the jackknife criterion was particularly encouraging. In Monte Carlo studies, the percentage of correct identification of true group membership for individual cases is infrequently used as a measure. Where it has been used, the results obtained suggest recovery between 60% and 80% for hierarchical analyses (Kuiper & Fisher, 1975; Mezzich, 1978). Upper limits for classification accuracy may be on the order of 95% for partitioning analyses (Mezzich, 1978). While the data base is too limited to provide a definitive frame of reference for the current findings, the present recovery rates compare favorably to these Monte Carlo results. Even using the 95% value as a reasonable estimate of the upper limit of recovery accuracy, the 85% agreement in the present findings would represent 89% of the maximum. Furthermore, comparison of the kappa and Rand values obtained in this study to the results obtained in Monte

Carlo studies (Milligan, 1981b; Milligan & Cooper, 1987) would lead to the same general conclusion. These values are more impressive if one assumes that the small group differences on Openness to Experience indicate that this dimension is irrelevant to the distinction between reactive and resistant individuals. If so, this variable basically introduces error into the assessment of intercase distances and should degrade classifications somewhat (Milligan, 1980). In perspective, the level of classification accuracy achieved in the present study was very good.

Additional support for the reality of the subtypes identified in these analyses was provided by the convergence between the data and Meehl's (1973) hypothetical model for typologies. This model is applicable to reactivity because it is derived from the assumption that key phenotypic behaviors are manifestations of genetic predispositions modulated by social history. Meehl's (1973) model predicts that reactivity indicators will be uncorrelated within groups of individuals comprised of a single type. While the five dimensions comprising the personality model employed in this study are conceptually independent, scale scores on these dimensions typically are correlated in the population at large (e.g., Costa & McCrae, 1985, 1989). However, the zero correlation criterion was met for 4 of the 5 personality dimensions considered in the present study when the computations were performed within the different groups. The exception was Openness to Experience, and this dimension may not be germane to the typology.

Although the preceding points support the position that reactive and resistant subtypes can be identified in the recruit population, it is important to note two differences between the hypothetical reactivity profile in Table 1 and the empirical profiles. The more substantial difference was that the empirically-defined stress reactive group was below average on Agreeableness, rather than above average as hypothesized. The difference between predicted and empirical results was the opposite for stress resistant individuals. This difference between the hypothetical and empirical profiles arises from an initial implicit equation of low Agreeableness with hostility and aggression and high Agreeableness with acquiescence in social interactions. In hindsight, the difference between reactive and resistant individuals may be more subtle than the simple presence or absence of aggressive, hostile behavior. As Sapolsky (1990a, 1990b) has noted, resistant animals apparently make instrumental use of aggression, employing it only when it will advance their position, while reactive animals are less discriminating in their use of aggression (Sapolsky, 1990a, 1990b). Also, dominance within a primate colony can be the

product of social coalitions. An individual's capacity to participate in such coalitions may depend on the ability to tolerate and trust other individuals. The present findings make sense if the instrumental use of aggression and coalition formation are characteristics of nonreactive individuals, while reactive individuals use aggressive behavior indiscriminately and lack the basic interpersonal trust to form the social ties necessary for effective coalitions.

The second important difference between the hypothetical reactivity profile and the empirical reactivity profile was the weakness of Openness to Experience as a discriminating factor between the two groups. Here again, hindsight suggests that the initial assignment could represent a mistaken generalization. The key to the discrepancy in this instance may be to distinguish between interest and overt behavior in new situations. Inhibited behavior in novel settings may indicate fearfulness, a component of neuroticism, more than a lack of intrinsic interest in exploring the situation. Also, a novel situation may be excessively stimulating for introverts until they have time to adjust to the situation (Eysenck, 1967, 1981). Thus, the inhibited behavior observed in children may be a product of other psychological factors than Openness to Experience.

The clusters defined in this paper constitute a provisional measurement model for testing hypotheses about the stress reactivity typology. These clusters were shown to be replicable, defined with sufficient precision to classify individuals with reasonable accuracy, to conform to one prediction based on Meehl's (1973) general model for psychobiological typologies, and to reasonably approximate predicted personality profiles and subtype frequencies estimated from other stress reactivity research. While these findings are encouraging, these clusters do not confirm the stress reactivity typology any more than the development of an internally consistent measure comprised of face valid items validates a psychological construct that assumes the existence of a continuum of differences in a psychological state or trait. The validity of these clusters as indicators of status with respect to the conceptual constructs comprising the stress reactivity typology must be established by further research to anchor these clusters in a network of empirical relationships that provides context for interpreting the findings. In this regard, the provisional clusters are directly comparable to any newly developed set of measures derived from a theory of individual differences.

The present findings provide sufficient reason to believe that additional research to use the provisional measurement model to evaluate the reactivity construct as a basis for understanding differences in the behavior of adult humans under stress is justified. The most important direction for such studies would be verification of fundamental stress reactivity assumptions. The assumption that stress reactivity is manifested in overlapping behavioral and endocrine responses under stress is one key topic. Tests of the presumed genetic basis for the typology would be another valuable addition. Demonstrating predictable relationships to behavior and mood in novel situations would be a third basis for evaluation of the construct. As noted in the introduction, the direct comparison of a reactivity typology with dimensional models should be a central concern in these evaluations. The possibility of refining the measurement model provided here should be one consideration in any such undertaking. Key issues for this component of the validation process include the utility of distinguishing instrumental aggression from general aggressiveness, the proper place of the Openness to Experience construct in the typology, and whether the typology applies to females as well as the males studied here. The provisional reactivity typology presented here provides a useful starting place for examining stress reactivity constructs as predictors of socially and personally significant differences in behavior in adult humans.

REFERENCES

Atkinson, J. W., & Feather, N. T. (Eds.) (1966). A theory of achievement motivation. NY: Wiley.

Blashfield, R. K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. Psychological Bulletin, 83, 377-388.

Blashfield, R. K., & Aldenderfer, M. S. (1988). The methods and problems of cluster analysis. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of multivariate experimental psychology 2nd Ed.) (pp. 447-474). NY: Plenum.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 70, 213-220.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd Ed.) (pp. 57-59). Hillsdale, NJ: Erlbaum.

Conley, J. J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. Personality and Individual Differences, 5, 11-25.

Costa, P. T., Jr., & McCrae, R. R. (1985). The NEO personality inventory manual. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. Journal of Personality and Social Psychology, 54, 853-863.

Costa, P. T., Jr., & McCrae, R. R. (1989). The NEO PI/FFI manual supplement. Odessa, FL: Psychological Assessment Resources.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. Annual Review of Psychology, 41, 417-440.

Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. Multivariate Behavioral Research, 16, 149-170.

Dunn, O. J. (1958). Estimation of the means of dependent variables. Annals of Mathematical Statistics, 29, 1095-1111.

Eysenck, H. J. (1967). The biological basis of personality. Springfield, IL: Thomas.

Eysenck, H. J. (1981). General features of the model. In H. J. Eysenck (Ed.), A model for personality (pp. 1-37). NY: Springer-Verlag.

Gangestad, S., & Snyder, M. (1985). "To carve nature at its joints": On the existence of discrete classes in personality. Psychological Review, 92, 317-349.

Higley, J. D., & Suomi, S. J. (1989). Temperament reactivity in non-human primates. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), Temperament in childhood. NY: Wiley.

Hoelter, J. W. (1983). The analysis of covariance strructures. Goodness-of-fit indices. Sociological Methods and Research, 11, 325-344.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), Personality -- current theory and research PP. 55-90). Nebraska Symposium on Motivation, 1982. Lincoln, NE: University of Nebraska Press.

Hubert, L., & Arabie, P. (1985). Comparing partitions. Journal of Classification, 2, 193-218.

Jemmott, III., J. B., Hellman, C., McClelland, D., Locke, S. E., Kraus, L., Williams, R. M., & Valeri, C. R. (1990). Motivational syndromes associated with natural killer cell activity. Journal of Behavioral Medicine, 13, 53-73.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin (Ed.), Handbook of personality: Theory and research (pp. 67-100). NY: Guilford.

Kagan, J. (1989). Temperamental contributions to social behavior. American Psychologist, 44, 668-674.

Kagan, J., Reznick, J.S., & Snidman, N. (1986). Temperamental inhibition in early childhood. In R. Plomin & J. Dunn (Eds.), The study of temperament: Changes, continuities, and challenges (pp. 53-67). Hillsdale, NJ: Erlbaum.

Kagan, J., Reznick, J. S., Snidman, N., Gibbons, J., & Johnson, M. O. (1988). Childhood derivatives of inhibition and lack of inhibition to the unfamiliar. Child Development, 59, 1580-1589.

Kagan, J., Reznick, J. S., & Gibbons, J. (1989). Inhibited and uninhibited types of children. Child Development, 60, 838-845.

Kuiper, F. K., & Fisher, L. A. (1975). A Monte Carlo comparison of six clustering procedures. Biometrics, 31, 777-783.

McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. Journal of Personality and Social Psychology, 38, 793-800.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. Journal of Personality and Social Psychology, 52, 81-90.

McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs type indicator from the perspective of the five-factor model of personality. Journal of Personality, 57, 17-40.

Meehl, P. E. (1973). MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. In P. E. Meehl, Psychodiagnosis: Selected papers (pp. 200-224). Minneapolis: University of Minnesota Press.

Mezzich, J. (1978). Evaluating clustering methods for psychiatric diagnosis. Biological Psychiatry, 13, 265-281.

Miller, R. G. (1974). The jackknife -- a review. Biometrika, 61, 1-17.

Milligan, G.W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika, 45, 325-342.

Milligan, G. W. (1981a). A Monte Carlo study of thirty internal criterion measures for cluster analysis. Psychometrika, 46, 187-199.

Milligan, G. W. (1981b). A review of Monte Carlo tests of cluster analysis. Multivariate Behavioral Research, 16, 379-407.

Milligan, G. W., & Cooper, M. C. (1986). A study of the comparability of external criteria for hierarchical cluster analysis. Multivariate Behavioral Research, 21, 441-458.

Milligan, G. W., & Cooper, M. C. (1987). Methodology review: Clustering methods. Applied Psychological Measurement, 11, 329-354.

Myers, I. B. (1980). Introduction to type. Palo Alto: Consulting Psychologists Press.

Plomin, R., Chipuer, H. M., & Loehlin, J. C. (1990). Behavioral genetics and personality. In L. Pervin (Ed.), Handbook of personality: Theory and research (pp. 225-243). NY: Guilford.

Plomin, R., & Dunn, J. (Eds.) (1986). The study of temperament: Changes, continuities and challenges. Hillsdale, NJ: Erlbaum.

Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association, 66, 846-850.

Sapolsky, R. M. (1990a). Stress in the wild. Scientific American, 262, 116-123.

Sapolsky, R. M. (1990b). Adrendocortical function, social rank, and personality among wild baboons. Biological Psychiatry, 28, 862-878.

Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. Journal of Personality and Social Psychology, 51, 125-139.

SPSS, Inc. (1988). SPSS-X user's guide (3rd Edition). Chicago: SPSS, Inc.

Strelau, J. (1983). Temperament, personality, activity. San Diego: Academic Press.

Strube, M. J. (1989). Evidence for the Type in Type A behavior: A taxometric approach. Journal of Personality and Social Psychology, 56, 972-987.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.

Appendix

Table A-1

Personality Profiles for 5-Cluster Solution

Cluster	N	E	O	A	C
R+ Mean	2.37	2.15	2.30	2.19	1.84
S.D.	.37	.36	.34	.38	.33
A- Mean	2.01	2.24	2.12	2.16	2.39
S.D.	.30	.28	.26	.33	.26
A+ Mean	1.87	2.55	2.52	2.59	2.37
S.D.	.33	.30	.32	.29	.32
C+ Mean	1.98	2.45	2.24	2.32	2.90
S.D.	.32	.47	.28	.30	.24
R- Mean	1.45	2.71	2.43	2.68	3.17
S.D.	.33	.31	.35	.35	.31

NOTE: N = Neuroticism, E = Extraversion, O = Openness to Experience, C = Conscientiousness, A = Agreeableness. R+ = Reactive Group (n = 549), A- = Low Agreeableness Group (n = 889), A+ = High Agreeableness Group (n = 711), C+ = High Conscientiousness Group (n = 643), and R- = Resistant Group (n = 536).

Table A-2

Linear Classification Function Coefficients for Group Assignment

Cluster	N	E	O	C	A	Constant
R+	20.01632	21.57863	20.77468	24.46163	20.91739	-117.9561
A-	16.71819	23.07574	19.29407	30.95208	20.67885	-123.9828
A+	15.15774	25.64784	23.51161	31.26669	24.67052	-147.1074
C+	16.17387	25.43409	20.66752	37.24076	22.35358	-151.8166
R-	11.03394	28.00863	23.15186	40.97130	25.70719	-175.1719

NOTE: See Table A-1 for explanation of row and column labels. These classification weights are appropriate only if the scale scores are computed as the average response to the items comprising the scales.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
26 AUG 91

3. REPORT TYPE AND DATE COVERED
Interim Mar 91 - Aug 91

4. TITLE AND SUBTITLE
Stress Reactivity: Five-Factor Representation of a Psychobiological Typology

5. FUNDING NUMBERS
Program Element: 61153N
Work Unit Number:
MR04101.00A-6004
ONR Reimbursable

6. AUTHOR(S)
Vickers, Jr., Ross R.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Health Research Center
P. O. Box 85122
San Diego, CA 92186-5122

8. PERFORMING ORGANIZATION
Report No. 91-26

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Medical Research and Development Command
National Naval Medical Center
Building 1, Tower12
Bethesda, MD 20889-5044

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A typology contrasting stress reactive individuals with stress resistant individuals has been proposed based on studies of children and nonhuman primates. The present study tested for these hypothesized types in U.S. Navy recruits who completed the NEO Personality Inventory. Hierarchical agglomerative and partitioning cluster analyses were conducted in 10 samples of U.S. Navy recruits (n = 331-335). Replicability of clusters across subsamples was a primary criterion for choosing between alternative cluster solutions. Applying this criterion, up to five clusters could be identified reliably. The five-factor personality profiles for two types in the five-cluster solution approximated a priori predictions about reactive and resistant profiles well enough to indicate a reactivity typology is a viable model of personality differences in this population.

14. SUBJECT TERMS
personality
typology
cluster analysis
stress
military recruits

15. NUMBER OF PAGES
37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102